ALIGNER popAI STARLIGHT

BUILDING A SECURE AND ETHICAL AI ECOSYSTEM FOR CIVIL SECURITY: THE SU-AI CLUSTER

FROM RESEARCH TO POLICY RECOMMENDATIONS

OCTOBER 2023

THE SU-AI CLUSTER

In 2021, the European Commission funded the first cluster of three projects under the H2020 secure societies call "SU (Security Union)-AI" that had a specific focus on artificial intelligence (AI) and civil security: ALIGNER - Artificial Intelligence Roadmap for Policing and Law Enforcement; popAI - A European Positive Sum Approach towards AI tools in support of Law Enforcement and safeguarding privacy and fundamental rights; and STARLIGHT - Sustainable Autonomy and Resilience for LEAs using AI against High priority Threats.

These three research projects form the nucleus of the European AI hub in support of law enforcement to enhance the AI competitiveness of Europe, while strengthening society's perception of the EU as an area of freedom, justice, and security.

The three projects were specifically funded to complement each other. While ALIGNER mainly focuses on the operative needs of police and law enforcement to derive research and policy recommendations, popAI focuses on the societal and ethical dimension of AI for policing and law enforcement. STARLIGHT, on the other hand, co-develops practical AI applications.

Since its inception, the SU-AI cluster has collaborated closely—not just within the cluster but also with other projects and initiatives beyond—to make use of synergies and ensure AI can be put into law enforcement practice in an ethical, societal, and legal way.

Within multiple joint activities—most notably an event co-organised with DG Home on the ethical and legal aspects of AI for law enforcement—to derive complementary approaches and joint recommendations.

This brochure gives an overview of the applied approaches and gathers the main cluster recommendations.

ABOUT US

ALIGNER

aligner-h2020.eu

ALIGNER brings together European actors at the cross-section of AI, law enforcement, and policing to collectively identify and discuss needs for paving the way for a more secure Europe in which AI supports law enforcement agencies while simultaneously empowering, benefiting, and protecting the public. To achieve this, ALIGNER has established a forum for exchange between practitioners, civil society, policymaking, research, and industry to design an AI research and policy roadmap meeting the operational, cooperative, and collaborative needs of police and law enforcement.

popAI

pop-ai.eu

popAI aims to increase awareness, boost trust, and encourage a constructive dialogue on the ethical use of AI in policing between EU policymakers, law enforcement agencies (LEAs), technology providers, and civil society. The popAI project follows a comprehensive, cross-disciplinary, and inclusive approach, mapping and engaging all ecosystem stakeholder groups through a series of knowledge-sharing modalities and collaboration tools. The project results include AI policy recommendations and multidisciplinary best practices, as well as an ethics toolbox for law enforcement agencies, organising the knowledge around AI within functionality and ethical and legal taxonomies. These practical tools are part of popAI's blueprint for a European AI hub for LEAs, a hub facilitating the exchange of knowledge between stakeholders and supporting the responsible use of AI by law enforcement.

STARLIGHT

Sustainable Autonomy and Resilience for LEAs using AI against High Priority Threats

starlight-h2020.eu

STARLIGHT aims to enhance the EU's strategic autonomy in artificial intelligence for law enforcement agencies. STARLIGHT achieves this by creating a community that brings together LEAs, researchers, industries, and practitioners within the security ecosystem to work together in a coordinated and strategic effort to bring AI into operational practices. Underpinning this approach are five key components that make up the STARLIGHT vision: (i) improve the understanding of AI across LEAs; (ii) provide opportunities for LEAs to exploit AI solutions; (iii) enable LEAs to protect AI systems; and (iv) combat the misuse of AI-supported crime and terrorism. The final component (v) focuses on boosting the uptake of AI across European LEAs while delivering long-term sustainability as part of the AI hub in support of law enforcement. A comprehensive analysis, understanding, and application of legal and ethical principles and the protection of fundamental rights across the entire AI development lifecycle support STARLIGHT's work.

ALIGNER, popAI, and STARLIGHT have received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement nos. 101020574, 101022001, and 101021797.

CO-DEVELOPING AI SOLUTIONS WITH LAW ENFORCEMENT PRACTITIONERS

STARLIGHT: SUSTAINABLE AUTONOMY AND RESILIENCE FOR LEAS USING AI AGAINST HIGH - PRIORITY THREATS

STARLIGHT, led by coordinator CEA (the French Atomic and Alternative Energies Commission), is a 50-partner-strong project designed to bolster the EU's strategic autonomy in artificial intelligence for law enforcement agencies. This is an ambitious four-year project, launched in 2021.

STARLIGHT brings together a consortium from 18 European countries, including 14 LEAs and Europol, from the security ecosystem under a coordinated and strategic effort to bring AI into LEAs' operational practices. The project takes a human-centric approach to designing and developing AI tools, carefully addressing potential ethical and legal implications, to deliver responsible, explainable, trustworthy, and accountable results aligned with EU principles and regulations. This report arrives at the midpoint of the STARLIGHT project and represents an excellent opportunity to take stock of the complementary outcomes of the AI cluster projects, reflect on the progress and approach to date, and highlight the forthcoming activities.

THE ROAD SO FAR

The initial stages of the project focused on the current landscape: gathering the views and experiences of LEAs with artificial intelligence, building the STARLIGHT community, and cataloguing existing and potential AI solutions under development. This left the project with a clear view of the capabilities, gaps, needs, and challenges for law enforcement in the context of AI, leading to a comprehensive set of user requirements and use case scenarios across the six core STARLIGHT domains: border security, counterterrorism, child sexual exploitation, cybercrime and cybersecurity, serious and organised crime, and the protection of public spaces. Concurrently, there was a strong focus on providing high-quality training and testing datasets for developing robust and accurate AI models.

STARLIGHT delivers new mechanisms for dataset creation and annotation, privacy-preserving measures for training and testing AI tools, and factors related to the legal and ethical aspects of generating and exchanging data. Using this data, STARLIGHT builds reliable, accountable, responsible, and transparent tools customised for LEAs' needs. It covers various technological needs, including machine learning, machine reasoning, computer vision, natural language processing, prediction, and AI-based cybersecurity.

STARLIGHT has laid the foundations for implementing trustworthy and accountable AI solutions that fully respect the fundamental rights of citizens, the current legal frameworks applicable to AI, and ethical principles. STARLIGHT's ethical and legal observatory closely monitors and supports technical development and provides guidance and support to deliver legally and ethically responsible solutions.

TOOLFESTS AND CO-DEVELOPMENT CYCLES

STARLIGHT breaks the mould in developing and testing AI solutions for LEAs. Building on the best practices approach from previous related European projects like ASGARD, STARLIGHT has launched a series of ToolFests to enable research and industry partners to test and demo early prototypes and solutions, gain feedback from LEAs, and identify new use cases and applications. The ToolFests provide a unique forum for accelerating AI technology development and are a launchpad for the co-development (co-dev) cycles. Based on LEA-defined priorities and the available AI tools, LEAs and technical partners are matched in the co-devs to develop highly specific solutions that address a particular LEA need.

These partners worked closely together for six months, iterating, improving, and testing solutions, enhancing their technology readiness level (TRL), and highlighting the potential of AI for LEAs. Overall, six co-dev cycles will run throughout the project, not only delivering technical solutions but also enhancing LEAs' understanding of the use of AI solutions and constructing stronger links between the different actors in the security ecosystem. Finally, STARLIGHT also promotes the proactive safeguarding of AI solutions adopted by LEAs by providing solutions that bolster cybersecurity operations, protect AI solutions against adversarial attacks, and enable intelligent detection of cybersecurity threats.

THE FUTURE OF STARLIGHT

The 2023 Security Research Event is a critical juncture for STARLIGHT, presenting the whole project together with its sibling projects (popAI and ALIGNER) and hosting a new ToolFest to demonstrate existing AI tools to LEAs. This event will also be followed by the first full pilot, a STARLIGHT internal event that brings together multiple AI solutions emerging from the co-development cycles and ToolFests. These tailored pilot scenarios address real operational needs, enabling the project to pivot and address additional domains and challenges relevant to LEAs and Europol. To ensure longevity and long-term impact, innovation uptake is prioritised through an innovative community-building and management process that gets to the core of LEA needs and the challenges for innovation uptake, as well as identifying where existing solutions can be adapted, integrated, or updated to meet the requirements of individual forces. This development pathway ensures that STARLIGHT is able to work at both European and national levels and deliver value across the security ecosystem. Meanwhile, future ToolFests, co-development cycles, and pilots will continue to demonstrate the latest advances in AI for addressing LEAs' needs.

STARLIGHT, working with ALIGNER and popAI, will forge a sustainable European AI community, unifying stakeholders from the security and AI domains. This community will drive the adoption of AI technologies in the daily operational activities of LEAs. The success of this endeavour depends on creating more effective and efficient collaborative networks that bridge EU agencies, law enforcement agencies, research institutions, small and medium enterprises, and other industry stakeholders. Such an approach, in conjunction with the co-development cycles, will fulfil STARLIGHT's ambition to elevate Europe's collective response to criminal threats and foster the emergence of a robust EU AI-based security sector.

DERIVING POLICY AND RESEARCH RECOMMENDATIONS

THE ALIGNER APPROACH

ALIGNER derived policy and research recommendations via regular exchanges with experts from its law enforcement agencies advisory board, scientific, industrial, and ethical advisory board, the SU-AI cluster, as well as other research projects and initiatives during three successive workshops, ensuring comprehensive insights and perspectives were considered in shaping these valuable recommendations.

The workshops focused on topics like 'AI as a Crime and Security Threat', 'AI in Service of Policy and Law Enforcement', and 'Capability requirements of police and law enforcement relating to AI technology'. The workshops were supported by an online survey to derive needs from practice and identify gaps in policy and research, as well as by a screening of literature and existing research results. The derived set of policy recommendations was then validated by the experts via the fourth ALIGNER workshop.

THE popAI APPROACH

popAI introduced an inclusive, holistic, and iterative approach to policy recommendations, (i) involving different stakeholders, including citizens and citizen representatives among others, thus ensuring all voices and perspectives are taken into consideration (inclusive), (ii) considering that AI technologies and their usage in the security domain can have implications at multiple levels and across different time frames and for different stakeholder segments (holistic), and (iii) facilitating an ongoing dialogue and a constantly updated approach to the provision of policy recommendations (iterative).
popAI prepared two annual policy briefs through evidence-based analysis, including practical recommendations based on multiple activities conducted in the project: policy labs, workshops, as well as crowdsourcing, and social listening tools. The novel popAI approach responds to the how and why of policy across the implementation and usage of trusted AI tools in the security domain by combining both (a) bottom-up (citizen and stakeholder-driven), and top-down (from research and consortium partners) recommendations; and (b) reactive (current state of affairs) and proactive recommendations (future state of affairs).

THE STARLIGHT APPROACH

STARLIGHT's methodology utilises a cyclical approach and emphasises co-design and co-creation, ensuring integration among industrial, research, and law enforcement partners, leading to well-designed, trustworthy, and operationally relevant AI solutions. By engaging all stakeholders throughout the development lifecycle, we ensure that our solutions align with the practical needs and ethical considerations of the industries and communities we serve.

RECOMMENDATIONS FROM THE SU-AI CLUSTER

A proposal for a regulation laying down harmonised rules on artificial intelligence, the Artificial Intelligence Act, was issued by the European Commission on April 21, 2021. The AI Act constitutes the world's first comprehensive AI regulatory framework, puts the human in the centre (human-centered approach), and is a horizontal regulation applicable to AI systems of all sectors that are classified according to the risk they pose to the affected individuals (risk-based approach).

Further to the text proposed by the European Commission, the Council of the European Union issued its General Approach on November 25, 2022, with the amendments being formally adopted by the Council on December 6, 2022. Most of the main issues of interest for security, migration, and borders are likely to be discussed in the forthcoming months.

The Directorate-General for Migration and Home Affairs (DG HOME) will be in touch with the Directorate-General for Communications Networks, Content and Technology (DG CNECT) to support the negotiations. The SU-AI cluster contributes to the process through the development of relevant AI policy recommendations and a series of joint policy-oriented events and workshops and will remain in close collaboration with the relevant DGs.

The SU-AI cluster has developed six policy recommendations. The table below provides a systematic overview of these recommendations. This overview adapts the policy ontology originally developed by popAI, identifying for each recommendation at what level (Societal, Regulatory, Organisational, or Research) a recommendation should be implemented, whether the recommendation is reactive, targeting the current state-of-play, or proactive, anticipating new policy actions, who the target audience for the recommendation is, and which themes or aims are addressed by the recommendation. The recommendations are then described in more detail in the remainder of the section.

RECOMMENDATION	IMPLEMENTATION LEVELS	REACTIVE PROACTIVE	TARGET AUDIENCES	THEMES/AIMS
Provide common guidelines and unbiased specialist support to LEAs for the development, procurement, deployment, and use of AI technology.	Regulatory, Organisational		EU Parliament, European Commission, Member State Parliaments, Ministries, LEAs	Fairness, Transparency, Equality, Privacy, Human Rights, Non-Discrimination, Minimize misuse, AI Applicability
Establish unified frameworks for the evaluation of AI tools during development and deployment ensuring their ethical, legal, and societal compliance.	Regulatory, Organisational, Research		EC DG HOME, EU Parliament, European Commission, Research Institutes, Industry, LEAs	Fairness, Transparency, Equality, Privacy, Human Rights, Non-Discrimination, Trustworthy AI

RECOMMENDATION	IMPLEMENTATION LEVELS	REACTIVE PROACTIVE	TARGET AUDIENCES	THEMES/AIMS
Review existing and establish new legal mechanisms to ensure that AI systems and their use are ethical, legal, and societally acceptable.	Regulatory		EU Parliament, European Commission, Member States Parliaments	Fairness, Transparency, Equality, Privacy, Human Rights, Non-Discrimination, Minimize misuse, Trustworthy AI, AI Applicability
Develop meaningful dialogue between regulators, LEAs, researchers, industry, and civil society organisations to strengthen citizens' confidence in the use of AI tools by LEAs.	Regulatory, Organisational, Research, Societal		Member States Parliaments, Ministries, LEAs, Research Institutes, Industry, Civil Society Organisations	Diversity, Transparency, Social Inclusion, Awareness, Trustworthy AI
Support and invest in the development of guidelines for gender-sensitive and gender-responsive policing in the AI era.	Regulatory, Organisational, Societal		EC DG HOME, Ministries, LEAs	Diversity, Equality, Social Inclusion
Extend and adapt European and national research programmes to better facilitate evidence-based, participatory research into LEA needs regarding AI, the potential implications of the use of AI by LEAs, and the potential criminal use of AI.	Regulatory, Research		European Commission, Ministries, National Funding Agencies, Research Institutes, and Civil Society Organisations	Social Inclusion, Trustworthy AI, AI Applicability

1ST RECOMMENDATION

PROVIDE COMMON GUIDELINES AND UNBIASED SPECIALIST SUPPORT TO LEAS FOR THE DEVELOPMENT, PROCUREMENT, DEPLOYMENT, AND USE OF AI TECHNOLOGY.

Interactions during multiple activities of the SU-AI Cluster, including exchanges with other projects, survey results, and other research activities [1], highlight the need for and lack of clear guidelines for police and law enforcement agencies regarding the development, procurement, deployment, and use of AI technologies.

This includes, first and foremost, guidance on the reliable evaluation of the ethical, legal, and societal implications of the use of AI, supporting the effectiveness of AI evaluations by moving away from a black box approach towards explainable AI, as well as target-group-specific training.

A specific issue in the development and deployment of AI relates to data protection and the necessary trade-off between protecting personal, sensitive data and the need for large 'real-world' datasets for training applicable AI models. Specific guidance on how to ensure data protection while simultaneously allowing for training AI models with real-world applicability is very much needed. However, guidelines alone will not be sufficient. The complex, dynamic, but at the same time highly regulated environment in which police and law enforcement agencies operate requires that they have access to unbiased, specialist support during the development, procurement, deployment, and use of AI technologies.

To achieve this, the EU and Member States should establish a European network of multidisciplinary, trustworthy AI support centres to support LEAs with choosing, procuring, and integrating AI technologies. On a European level, Europol and its EU Innovation Hub for Internal Security [2] might be the prime target to establish such a centre where LEAs can safely test and evaluate AI technologies in clearly defined 'sandboxes'.

However, this support centre needs to be complemented by national centres to lower hurdles for engagement (e.g., due to language barriers). Such centres need to be independent entities, funded nationally and not dependent on other funding mechanisms, that can then provide a form of external certification for AI technologies, also covering algorithm audits and evaluations of the extent to which systems use 'democratic' data in addition to 'robust' algorithms.

Critically, these support centres should also act as societal nodes where different actors affected by AI technologies (i.e., civil society organisations) as well as specialists in ethics, law, and AI development engage in discussions with police and law enforcement agencies on whether, how, and when to employ which AI technology (see also recommendation 4). For this reason, and to provide a neutral testing ground, these support centres should explicitly not develop AI technologies themselves.

Without such guidance and support, there is a high risk of abuse and/or misuse of AI technologies, leading to stigmatisation, discrimination, and potential violence against privacy and human rights. As such, it is important that the EU and Member States encourage and support the development of clear guidelines and support structures for the use of AI technologies by LEAs.

ESTABLISH UNIFIED FRAMEWORKS FOR THE EVALUATION OF AI TOOLS DURING DEVELOPMENT AND DEPLOYMENT, ENSURING THEIR ETHICAL, LEGAL, AND SOCIETAL COMPLIANCE.

The guidelines and support needed to ensure ethical, legal, and societal compliance, as well as the actual applicability of AI technologies, must be grounded in evidence-based, unified evaluation frameworks. Given the special role of LEAs within society, such assessment frameworks will need to follow a broader approach to impact assessment.

As identified by popAI, the literature proposes several AI tool assessment frameworks [3, 4, 5, 6], as well as methods that provide indicators of risks a company might face when adopting an AI tool, while also including mitigation actions and best practices that might be followed. Each of these frameworks includes different guidelines, assessment criteria, and mitigation recommendations concerning the adoption of AI.

However, most of them focus on the private sector, resulting in a lack of assessment frameworks and clear implementation procedures that provide guidelines, recommendations, and mitigation indicators for the adoption of AI tools in the public sector (see also recommendation 1). The AP4AI framework for assessing the accountability of AI systems as well as the ALIGNER fundamental rights impact assessment [7] (which is based on the MAGNETO ethical risk assessment form) take steps in this direction but need to be further aligned with other frameworks.

Therefore, there is an ongoing need for more extensive research, both on the development of such frameworks and the development of the corresponding interdisciplinary assessment measures and metrics.

With such frameworks, the adoption of an AI tool can be evaluated against a set of interdisciplinary metrics, developed in an inclusive manner, including the system scope, performance, usability, data used for training and evaluation including ethical processing, human rights impact, as well as ensuring compliance with data protection. Such frameworks should also include specific guidelines on mitigating bias of AI models and datasets.

REVIEW EXISTING AND ESTABLISH NEW LEGAL MECHANISMS TO ENSURE THAT AI SYSTEMS AND THEIR USE ARE ETHICAL, LEGAL, AND SOCIETALLY ACCEPTABLE.

Operative guidelines for the development, procurement, deployment, and use of AI technologies, based on evidence-based, unified evaluation frameworks, will need to be flanked by binding legal mechanisms to ensure that these technologies are ethical, legal, and societally acceptable.

The EU AI Act is a step in this direction, although based on numerous discussions with representatives from LEAs, civil society, research, industry, and policy, there remain valid concerns from different actors on its definition of AI (too broad), the exemptions included for high-risk AI technologies (too many), and its effect when put into place (too bureaucratic).

A valid approach to alleviating these concerns might be the development of a LEA-specific AI directive (similar to the Law Enforcement Directive [8]). Regardless of these issues, any legal mechanism on the EU and national level related to the use of AI technologies by LEAs needs to ensure that there is always a competent and knowledgeable 'human in the loop' if AI technology is used in critical decision-making processes.

The nature of the work carried out by LEAs and its impact on individuals and on wider society require that AI technology not directly replace human decision-making. Without this safeguard, all the checks and balances that are intrinsic to decision-making in LEAs cannot occur or are compromised, e.g., the fundamental issues of accountability, explicability, transparency, and compliance with the rule of law.

Even if an AI technology does not directly take decisions but only informs a human operator, the information provided via the AI technology has the potential to influence the decision. As such, it becomes of utmost importance that the data and information on which AI technology is trained, tested, validated, and used is accurate and does not perpetuate existing biases and stereotypes present in society.

Legal mechanisms in the EU and nationally should support the continuous, inclusive, and multidisciplinary monitoring of AI technology across its lifecycle. In particular, EU Member States should invite civil society organisations and create joint working groups that will check the individual AI technologies used by LEAs to highlight potential issues from such usage (a posteriori monitoring and assessment). These joint working groups should also be consulted when designing and developing new AI technologies that will be applied in the future (a priori monitoring and assessment). The purpose is to improve and adapt these technologies appropriately to ensure that they protect citizens' rights. This will support the use of existing technologies as well as the development of new ones to meet current needs.

The interaction between different actors related to the use of AI technologies by police and law enforcement agencies should be continuous (e.g., via the AI support centres suggested in recommendation 2) and should strengthen the involvement of civil society in all stages of the operation of an AI technology (design, implementation, maintenance, and upgrade).

To facilitate this interaction, the European Commission and EU Member States need to better promote and ensure citizens' awareness regarding the existence and implementation of AI technology and enable objections to potential unjust decisions.

Open discussions between different actors related to the use of AI technologies by LEAs can support transparency at every stage to minimise the risks of discrimination. In addition, this should also be considered in the procurement of systems, where, for example, the technical specifications must be accepted by civil society organizations and agencies, while monitoring and assessment by representatives of social and other bodies should be foreseen in the system implementation phase.

4TH RECOMMENDATION

DEVELOP MEANINGFUL DIALOGUE BETWEEN REGULATORS, LEAS, RESEARCHERS, INDUSTRY, AND CIVIL SOCIETY ORGANISATIONS TO STRENGTHEN CITIZENS' CONFIDENCE IN THE USE OF AI TOOLS BY LEAS.

Civil society organisations are often not included in consultations regarding the employment of AI by LEAs.

Therefore, they express their concerns about emerging risks through announcements and legal actions. This gap is creating tensions that are constantly widening and damaging trust between the involved parties.

To repair trust issues, civil society organisations should be involved in open dialogues with European and national regulators, LEAs, researchers, and industries regarding the employment of AI technologies.

The results of such activities would enable European Member States to integrate European regulations (see recommendation 3) into their law, tailoring it to the culture and the specificities that govern their societies.

Civil society organisations should be actively involved in the process of designing and implementing AI technologies, as well as in the monitoring of the existing ones. They should also determine the best way to operate these systems to ensure human rights and generate acceptance among citizens.

5TH RECOMMENDATION

SUPPORT AND INVEST IN THE DEVELOPMENT OF GUIDELINES FOR GENDER-SENSITIVE AND GENDER-RESPONSIVE POLICING IN THE AI ERA.

This recommendation aims at the development of corresponding guidelines for the promotion of gender-sensitive and gender-responsive policing [9, 10, 11], especially in the era of AI. In 2010, the Women Police Officers Network (WPON) [12] was established with the support of the Southeast Europe Police Chiefs Association. Its scope was to place gender-sensitive policing at the top of the agenda of police reform and to serve as a platform for knowledge and experience exchange across police services and the needs and priorities of policewomen.

The network has so far achieved gender-sensitive policing with an emphasis on recruitment, selection, and professional development of women in police services. However, apart from this initiative, it is important in today's developed society to promote and develop appropriate actions and guidelines on the equality of all people in society to ensure no group is disadvantaged over another in its treatment by the police. [13]

This policy recommendation aims at the development of the corresponding guidelines from the EU and the relevant EU-funded projects to raise awareness about the position of women in police services and the development and implementation of sustainable solutions for the improvement of recruitment and retention of women personnel and their active involvement in the design and development of AI systems for security purposes.

In addition to gender-sensitive policing, the aim is to achieve gender-responsive policing, which means taking into account “the needs of all parts of the community, women and girls, men and boys, including minority or marginalised groups […] to ensure no group is disadvantaged over another in its treatment by the police”. [14]

To achieve both, the suggested guidelines should focus on the empowerment of gender equality in law enforcement with an emphasis on the needs of all parts of the community and facilitate the inclusive design and development of the corresponding AI technologies to ensure that no group is mistreated by the police.

Furthermore, these guidelines shall be based on the outcomes of the WPON and the Southeast Europe Police Chiefs Association, which proved that the absence of data leads to ineffective policies and legal frameworks and that it is necessary to include the appropriate information so that gender-sensitive policing can be enhanced.

6TH RECOMMENDATION

EXTEND AND ADAPT EUROPEAN AND NATIONAL RESEARCH PROGRAMMES TO BETTER FACILITATE EVIDENCE-BASED, PARTICIPATORY RESEARCH INTO LEAS NEEDS REGARDING AI, THE POTENTIAL IMPLICATIONS OF THE USE OF AI BY LEAS, AND THE POTENTIAL CRIMINAL USE OF AI.

EU and nationally funded security projects, and specifically those developing AI-driven technologies, have often raised concerns; see, for example, the FP7 project INDECT, Intelligent information system supporting observation, searching, and detection for the security of citizens in urban environment [15], which sparked concerns among Members of the European Parliament, calling on the European Commission to clarify its purpose [16].

The sometimes overly restrictive secrecy of such projects and the lack of publicly available information, together with the perceived potentially negative impact on civil liberties and fundamental rights, call for new approaches towards accountability.

One way to address these issues while maintaining the required level of security would be the establishment of specialised interdisciplinary ethics and legal committees that review proposals and ongoing research projects in the security domain on a continuous basis, so as to prevent potentially serious ethical, societal, and legal issues as well as abuse of human rights.

Aligned with recommendations 1 and 2, these committees should have ethical, legal, technical, organisational, and practical capabilities to assess an AI technology's ethical, legal, and societal compliance. This could act as a form of internal certification for research projects in relation to an AI technology's accountability and the ethical, inclusive, and secure-by-design AI systems in the course of research and development.

In addition, research conducted in the context of the H2020 project popAI identified the stakeholder groups involved in the research, development, use, and implementation of AI technology, as well as those who promote awareness regarding emerging risks and push for relevant policies.

These different categories of stakeholders should not be seen as 'rivals' but rather as key components of a unified ecosystem that co-shapes the development and use of AI in the security domain.

The identified stakeholders are, namely, LEAs, social and humanities researchers, policymakers, government and public bodies, technologists and data scientists, civil society organisations, national and local authorities, Information and Communication Technology (ICT) and software companies, and police academies.

Mapping EU-funded projects in the security domain, 348 different stakeholders were identified with the majority being ICT and software companies, followed by universities and research organisations.

It is recommended that the EC explore ways (i.e., call requirements, specifications) for EU-funded projects to include civil society organisations in the early stages of AI technology design and development, as they are underrepresented in the project consortium, and their voices are very important to preserving privacy and human rights.

Likewise, project partners were geographically mapped. The analysis indicated that various European countries, such as Albania, Denmark, and Ukraine, have been underrepresented to date in EU-funded projects in the security domain. Involvement of partners from underrepresented Member States would enable the inclusion of potentially cultural and geographic differences regarding the needs and acceptance of AI systems. Thus, it is recommended that the EC explore ways (i.e., call requirements, specifications) for EU-funded projects to include underrepresented Member States in AI design and development. [13]

Lastly, the implementation of recommendations 1-5 needs to be supported by further AI-specific research in the security domain.
This includes the development of guidelines aligned with the needs of LEAs (recommendation 1), assessment frameworks (recommendation 2), an evaluation of the existing legal mechanisms as well as their effects on LEA work (recommendation 3), stakeholder engagement techniques in the context of AI technologies for LEAs (recommendation 4), as well as guidelines for gender-sensitive and gender-responsive policing (recommendation 5). This also includes additional research into countering criminal use of AI technologies and employing AI technologies in support of LEAs in an ethical, legal, and societally acceptable way.

REFERENCES

[1] Laufs, J. and Borrion, H. (2021) "Technological innovation in policing and crime prevention: Practitioner perspectives from London," International Journal of Police Science & Management, pp. 1–20.

[2] Europol Innovation Lab (n.d.) "EU Innovation Hub for Internal Security." Available at: https://www.europol.europa.eu/operations-services-innovation/innovation-lab/eu-innovation-hub-for-internal-security (Accessed: September 2023).

[3] High-Level Expert Group (HLEG) (n.d.) "Assessment List for Trustworthy Artificial Intelligence (ALTAI)." Available at: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment (Accessed: September 2023).

[4] World Economic Forum (WEF) (n.d.) "AI Governance Framework." Available at: https://www.weforum.org/projects/model-ai-governance-framework (Accessed: September 2023).

[5] NOREA Guiding Principles (2021) "Trustworthy AI Investigation." Available at: https://www.norea.nl/uploads/bfile/a344c98a-e334-4cf8-87c4-1b45da3d9bc1 (Accessed: September 2023).

[6] Fraunhofer IAIS (n.d.) "AI Assessment Catalogue of Fraunhofer IAIS." Available at: https://www.iais.fraunhofer.de/en/research/artificial-intelligence/ai-assessment-catalog.html (Accessed: September 2023).

[7] Casaburo, D. and Marsh, I. (2023) "ALIGNER D4.2 Methods and Guidelines for Ethical and Law Assessment," H2020 ALIGNER, GA no. 101020574.

[8] European Parliament and Council of the European Union (2016) "Directive 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purpose of the prevention, investigation, detection, or prosecution of criminal offences," May 4. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32016L0680 (Accessed: March 2023).

[9] Clutterbuck, L. (2022) "ALIGNER D2.2 Archetypical Scenarios and Their Structure," H2020 ALIGNER, GA no. 101020574.

[10] United Nations Women 2021, Handbook on Gender-responsive Police Services for Women and Girls Subject to Violence, PDF, July 2021, 528 pages, English, ISBN (PDF): 9789216040710, DOI: https://doi.org/10.18356/9789216040710.

[11] Bonkat-Jonathan, L. and Ejalonibu, G. L. (2021) "A Review of Some Discriminatory Laws Against Women and the Need for Legislative-Gender Responsive Actions in Nigeria."

[12] Kekić, D., Đukanović, D., & Tomić, M., Women Police Officers Network (WPON).

[13] Ziouvelou et al. (2022) "popAI Policy Brief - 1st Year," Deliverable D1.6, H2020 popAI, GA no. 101022001.

[14] International Association of Women Police (n.d.) "Gender-Responsive Policing." Available at: https://www.iawp.org/Gender-Responsive-Policing-GRP (Accessed: September 2023).

[15] Cordis (n.d.) "INDECT (Intelligent Information System Supporting Observation, Searching, and Detection for the Security of Citizens in Urban Environments)." Available at: https://cordis.europa.eu/project/id/218086 (Accessed: September 2023).

[16] Euractiv (2011) "MEPs question 'Big Brother' urban observation project."

GET IN TOUCH

ALIGNER

aligner-h2020.eu

aligner-h2020

info@aligner-h2020.eu

---

popAI

pop-ai.eu

popai-project

info@pop-ai.eu

STARLIGHT

starlight-h2020.eu

starlight_h2020

starlight@cea.fr